DBSJ Journal, Vol.8, No.1
June 2009

Rank Optimization of Personalized Search

Lin LI♥ Zhenglu YANG♦
Masaru KITSUREGAWA♠

Augmenting the global ranking based on the linkage structure of the Web is one of the popular approaches in data engineering community today for enhancing the search and ranking quality of Web information systems. This is typically done through automated learning of user interests and re-ranking of search results through semantic based personalization. In this paper, we propose a query context window (QCW) based framework for Selective uTilization of search history in personalized leArning and re-Ranking (STAR). We conduct extensive experiments to compare our STAR approach with the popular directory-based search methods (e.g., Google Directory search) and the general model of most existing re-ranking schemes of personalized search. Our experimental results show that the proposed STAR framework can effectively capture user-specific query-dependent personalization and improve the accuracy of personalized search over existing approaches.

1. Introduction

Encoding human search experiences and personalizing the search result delivery through ranking optimization is a popular approach to enhance Web search. Although the general Web search today is still performed and delivered predominantly through search algorithms, e.g., Google's PageRank [17] based query independent ranking algorithms, the interests in improving global notion of importance in ranking search results by creating personalized view of importance have been growing over the recent years. We categorize the research efforts on personalized search into three classes of strategies: 1) query modification or augmentation [3], [26], 2) link-based score personalization [8], [9], [15], [17], [19], [22], and 3) search result re-ranking [4], [5], [12], [14], [26], [29], [30]. A general process of re-ranking is to devise efficient mechanisms to re-order the search result ranking using the global importance by personalized ranking criteria. Such criteria are typically derived from the modeling of users' search behavior and interests.

In this paper, we develop a rank optimization framework that promotes Selective uTilization of search history for personalized leArning and re-Ranking (STAR). Our STAR framework consists of three design principles and a suite of algorithms for learning and encoding user's short-term and long-term search interests and re-ranking of search results through a careful combination of recent and previous search histories. We show that even though short-term interests based personalization using the most recent search histories may be effective at times [15], [25], [26], it is generally unstable and fails to capture the changing behavior of the users. Furthermore, most of existing long-term interests based personalization using the entire recent and previous search histories fails to distinguish the relevant search history from irrelevant search history [4], [18], [30], making it harder to be an effective measure alone for search personalization.

Bearing in mind of these observations, our STAR framework advocates three design principles for rank optimization. First, we devise a so-called query context window (QCW) model to capture the user's search behavior through a collection of her per-query based click-through data. Second, we develop a query-to-query similarity model to distinguish the relevant search memories of personalized search behavior from irrelevant ones in the QCW of each user, reducing the noises incurred by using either a recent fragment or the entire QCW. Third, we develop a fading memory based weight function to carefully combine the frequency of relevant search behavior (long term interests) with the most recent search behavior (short term interests). To show the effectiveness of our STAR framework in quality enhancement of personalized search, we propose length and depth based hierarchical semantic similarity metrics and compare the effectiveness of four re-ranking strategies: 1) naive re-ranking that is query and time independent; 2) relevant search memory based re-ranking that is query dependent but time independent; 3) fading memory based re-ranking that is time dependent but query independent; and 4) hybrid re-ranking that is both query and time dependent. Our experiments show that the hybrid re-ranking scheme can effectively combine the previous and recent memories through a smooth and gradually fading memory based weighting function. More importantly, our experimental results show that the proposed STAR framework for personalized search and re-ranking can effectively capture user-specific query-dependent personalization preference and improve the accuracy of personalized search over the popular directory-based search methods (e.g., Google Directory search) and the general model of most existing re-ranking schemes of personalized search.

The remainder of this paper is organized as follows. The overview of our STAR framework is presented in Section 2. Then, we discuss building QCW-based user profiles and designing re-rank strategies in Section 3 and Section 4 respectively. Experimental results will be given in Section 5. Related works are reviewed in Section 6. Finally, we conclude the paper in Section 7.

♥ Dept. of Information and Communication Engineering, University of Tokyo
lilin@tkl.iis.u-tokyo.ac.jp
♦ Institute of Industrial Science, University of Tokyo
yangzl@tkl.iis.u-tokyo.ac.jp
♠ Regular Member Institute of Industrial Science, University of Tokyo
kitsure@tkl.iis.u-tokyo.ac.jp

2. The STAR Framework Overview

The goal of the STAR framework is to design a semantic rich user profile model to capture the query context and the search behavior of each user and intelligently utilize such user profiles to enhance the quality of personalized search by effectively re-ranking of the search results returned from a general purpose search engine. Figure 1 gives a sketch of the STAR framework, consisting of three main components.

The first component is the text classification module that performs hierarchical Web page classification. The popular way is to classify the documents into a directory-based ontology, such as Yahoo! Directory [11], ODP (http://dmoz.org) [27], and so on. Studies [1], [10], [16] preferred to building their own ontology. Thanks for the fact that hierarchical text classification is well studied in the field of text processing, in the first prototype design of our STAR framework we directly utilize the classified search results from Google Directory search.

The second component is the context aware learning of user's search behavior. We utilize the per-query based click-through data to capture query dependent context and search behavior and develop the query context window (QCW) model to encode such leaning process. By automatically generating QCW based user query profiles, the user learning module automatically captures the query dependent context of user search behavior. For example, our approach focuses on the user's visited search results (Web pages) which supply us with not only what kind of content a user is interested in (topics) but also how much the user is interested in them (click frequency).

The third component is the query and time dependent, hybrid re-ranking scheme that produces a new user-centric, query dependent rank list for each user query through three step process. First, it selects the relevant click records from the entire QCW of a user through the query-to-query similarity analysis. Second, it combines the recent search memories with the previous search memories through applying a fading memory based weighting function over the selected QCW click records of a user. Finally, it employs hierarchical semantic similarity measures to compute the personalized ranking of the search results returned from a general search engine. In the subsequent sections we will focus on the technical detail of the user learning module and the re-rank module.

3. QCW Based User Learning Module

Our STAR framework devises the query context window (QCW) to encode the user specific and query dependent search behavior. Given a user, her query context window consists of m query-dependent context click records, denoted as $u_1, u_2, \ldots, u_m$. Each click record in the QCW is composed of the submitted query, the topics associated with the click search results, the click frequency of each topic, and the returned search results of the given query. The topics are extracted from Google Directory, structured as a hierarchical tree, so that each click record has its own tree. This topic tree records the click behavior of a user on a specific query, which can tell us what kind technical details of click record selection are in the next

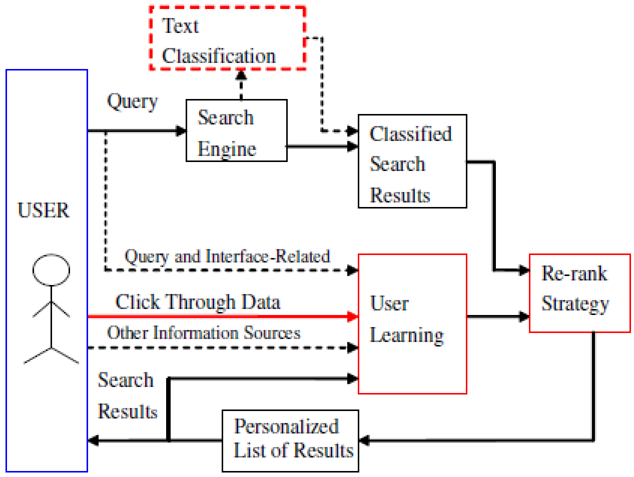


Figure 1 Overview of the STAR framework

section. Moreover, we implement each QCW as a queue. The tail of the queue holds most recently requested queries, while its head holds the least recently requested queries. When a new query is submitted, the corresponding record is added to the tail of the queue and the user model (QCW) is updated accordingly. This queue keeps the chronological order of different click records, which can easily differentiate the recent and old search histories for re-ranking strategies.

Figure 2 shows an example of QCW with three context records, each corresponds to one query and its context encoding of the query dependent click-through data. For example, a user inputs a query "Disneyland" to Google Directory search engine, and then input query "Disneyland" as a root node followed by the clicked topics. The search results are kept in the SRB. Node F is represented by the [Theme\Parks, 6] which means the user has clicked some search results associated with the topic "Theme\Parks" six times in this search. In addition, for each topic, we store the top four depth of its full path in Google Directory in a record. For example, the node F is actually stored as the [\Recreation\ThemeParks].

4. The Re-rank Module

The QCW based re-ranking module needs to address three key challenges: (1) how to select relevant context records from the entire QCW given a user query (Section 4.1); (2) whether all the selected query-relevant context records play the same role in re-ranking the search results of the current query (Section 4.2); (3) how should we re-order the search results (Section 4.3 and 4.4)? We use calligraphic upper-case alphabets to represent sets. The elements of a set are denoted by lower case alphabets. For example, U is the set of click records in QCW and u_i is an element (click record) of U. |U| is the cardinality of the set U.

4. 1 Selecting Relevant Click Records

Given a new input query, we first select the relevant QCW click records where the encoded queries are similar to the current input query by using a query-to-query similarity measure. Estimating the similarity (relatedness) between

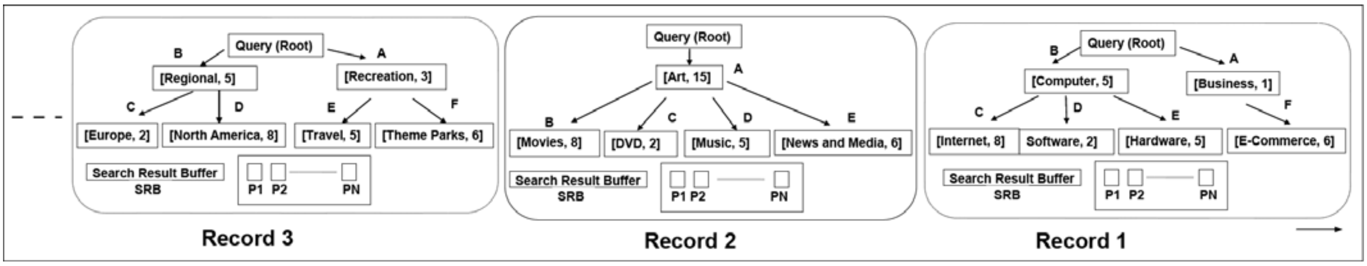


Figure 2 Query Context Window: click records are queued up in a chronological order

queries has a long history in traditional Information Retrieval [6], [21], [32]. It is still hot and active in various topics of Web Information Retrieval [2], [7], [31]. Up to now it has not been possible to prove that any of these measures outperforms all others in a large set of experiments [33].

The similarity between the two queries can be induced from the overlap of the two lists of search results (URLs) returned. Clearly, the query-result-vectors present a better similarity metric than query term-vectors [21]. As thus, we formally define the query-to-query similarity measure as follows:

$$Q(q^{u_i}, q_{in}) = \frac{SRB^{u_i} \cap P^{q_{in}}}{SRB^{u_i} \cup p^{q_{in}}} \qquad (1)$$

We can get the URL set of search results of q^{u_i} from the search result buffer SRB^{u_i} of the click record u_i, and the URL set of search results $P^{q_{in}}$ of q_{in} from the current search. The similarity between the two queries is estimated to the fraction of the intersection of the two URL sets. In our experiments, URL similarity is measured by their host name. We would like to note that our STAR framework can easily incorporate other similarity and specificity measures.

4. 2 Weighing Relevant Click Records

The selected click records are the collection of user's previous and recent search behaviors which reflect her interests. We assume that the user's interests will gradually decay as time goes on, so we assign more weights to more recent QCW click records and decreasing weights to older QCW click records to further improve the accuracy of the personalized search using a fading memory based weight function, defined as follows:

$$F(u_i) = e^{-\frac{\log 2}{hf \cdot |U|} \cdot (|U| - i)} \qquad (2)$$

where hf is a half fading parameter. In our experiments, hf is set in the range [0.1, 1]. After the click record u_i is selected as relevant according to Equation 1, its effect on the quality of personalized search depends on its temporal order. With increasing the value of hf, the rate of fading becomes slow and the weights on previous memories increase. This fading memory function unifies the user's long-term and short-term interests encoded in the QCW click records by assigning different weights to these click records appearing in different temporal order.

4. 3 Capturing Search Interests

After the relevant QCW click records and their weights are determined by Equation 1 and 2, the topics in these QCW click records are reflecting the user's current search interests. Now we can devise a re-ranking mechanism to re-order the search results by putting those that are more similar to the selected topics closer to the top of the final re-ordered rank list. In our STAR framework, the topics in the relevant QCW click records are structured in a semantic concept hierarchy as shown in Figure 2. Hierarchical similarity measures can be used to assess the similarity between the related topics and the search results of the given query. Let h be the depth of the subsumer (the deepest node common to two nodes), l be the shortest path length between two topics, and M be the maximum depth of topic directory possessed by a QCW click record.

Two combinations of depth and length based similarity measure defined as follows:

$$c1 \equiv HS(t_j^{u_i}, p_k^{q_{in}}) = \frac{2 \cdot h}{1 + 2 \cdot h} \qquad (3)$$

$$c2 \equiv HS((t_j^{u_i}, p_k^{q_{in}}) = e^{-0.2 \cdot l} \cdot \frac{e^{0.6 \cdot h} - e^{-0.6 \cdot h}}{e^{0.6 \cdot h} + e^{-0.6 \cdot h}} \qquad (4)$$

Equation 3 is a simple linear transformation function of the length and the depth, while Equation 4 transfers the length and the depth by a nonlinear function and then combines them by multiplication [13].

A QCW click record may record more than one topic depending a user's click behavior. We further define the similarity between a QCW click record u_i and a search result $p_k^{q_{in}}$ as:

$$s(u_i, p_k^{q_{in}}) = \frac{1}{|T^{u_i}|} \sum_{t_j^{u_i} \in T^{u_i}} \frac{HS(t_j^{u_i} \cdot p_k^{q_{in}}) \cdot c_j^{u_i}}{\sum_{c_j^{u_i} \in C^{u_i}} c_j^{u_i}} \qquad (5)$$

where each topic $t_j^{u_i}$ is weighted by its corresponding $c_j^{u_i}$ using the click frequency of the topic j. We further normalize the sum of hierarchical similarity scores through dividing it by the number of topics stored in a click record.

4. 4 QCW Based Re-ranking

In this section we will describe four strategies to re-order the search results of her current query.

1) Strategy 1 – Query and time independent scheme

$$s_1(U, p_k^{q_{in}}) = \frac{1}{|U|}\sum_{u_i \in U} s(u_i, p_k^{q_{in}}) \qquad (6)$$

"Strategy 1" is query and time independent, a naive strategy. Click records of different past queries are assigned equal weights regardless of the current input query. The similarity scores of past queries with a search result are summed together and divided by the number of click records (|U|) in U. "Strategy 1" thinks all the past histories (click records) are related to a user's current query. As we discussed in Section 1, the entire QCW includes noisy memories unrelated to the current query. Most of re-ranking based Web search personalization methods in the literature [4], [5], [12], [14], [18], [26], [27], [30]. "Strategy 1" can represent the general idea of these methods, compared with the following three strategies.

2) Strategy 2 – Query dependent scheme

$$s_2(U, p_k^{q_{in}}) = \frac{1}{|U|}\sum_{u_i \in U} Q(q^{u_i}, q_{in}) \cdot s(u_i, p_k^{q_{in}}) \qquad (7)$$

We define the "Strategy 2" as a query dependent and time independent strategy, which is selective about by using Equation 1 to weight these click records. Tan et al. [29] did preliminary discussion on query-dependent selection of user profile. However, their work is in the context of only exploiting long-term search histories of users and ignores the changes of user's interests with time.

3) Strategy 3 – Time dependent scheme

$$s_3(U, p_k^{q_{in}}) = \frac{1}{|U|}\sum_{u_i \in U} F(u_i) \cdot s(u_i, p_k^{q_{in}}) \qquad (8)$$

"Strategy 3" strengthens recent memories and weakens the effect of previous memories by applying Equation 2 to each QCW click record without the selection of relevant contexts in terms of the input query like "Strategy 2". If hf is set to a very small value, the previous memories cannot have an influential effect on re-ranking. Researches [15], [25], [26] emphasize that the most recent search is most directly close to the user's current information need, which can be regarded as a special case where hf is close to zero in "Strategy 3".

4) Strategy 4 – Query and time dependent scheme

$$s_1(U, p_k^{q_{in}}) = \frac{1}{|U|}\sum_{u_i \in U} F(u_i) \cdot Q(q^{u_i}, q_{in}) \cdot s(u_i, p_k^{q_{in}}) \qquad (9)$$

"Strategy 4" is query and time dependent, a hybrid strategy. As we know, users have their own characteristics of search behavior. To handle the most general case where we have many kinds of Web users and users will how different search behaviors, "Strategy 4" is designed to select relevant click records by Equation 1, but also assign greater weights to the more recent click records Equation 2.

Given one of the four strategies, a new relevant score will be calculated for each of search results. We output the list of the search results in order of their assigned scores. In the following experiments, we will evaluate the effectiveness of the above four re-rank strategies.

5. Experiments

5. 1 Experiment Setup and Evaluation Measure

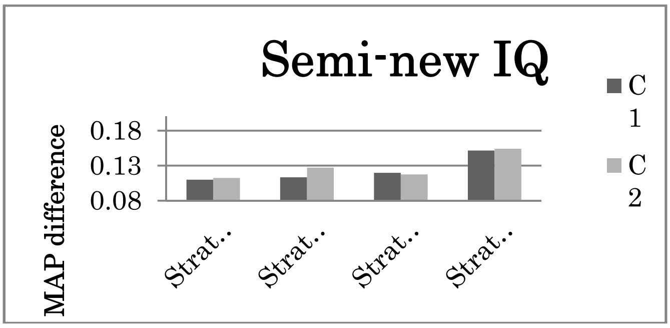
The goal of this paper is to achieve a personalized ranking by scoring the similarity between a user profile and the returned search results. Instead of creating our own Web search engine, we retrieve results from Google Directory search engine and use them as a baseline in the following evaluation. Moreover, as discussed in [23], informational queries (IQ) are such queries where the user does not have a special page in mind and intends to find out Web pages related to a topic. We further classified the goal of IQ into three categories: new IQ, semi-new IQ, and repeated IQ. A query is a new IQ if a user never searches such a topic before. It means that we cannot get the relevant search histories. A semi-new IQ has similar topical contents with some of the user's search histories. A repeated IQ refers to the query by which the user has already obtained the desired information, and is searching for it again. The following experiments will evaluate the performances of the semi-new and repeated IQs since our STAR framework wants to use the previous relevant memories to enhance the current search. For new IQs, collaborative information retrieval will be an interesting direction in our future work.

The evaluation of our framework is a challenge because currently there are no suitable query log data sets as a public benchmark. We created our own real dataset [12] which was collected over a ten-day period (From October 23rd, 2006, to November 1st, 2006). Twelve users are invited to search through our framework and judge whether the clicked results are relevant or not. Users were asked to input search queries related to their professional knowledge in the first four days, and search queries related to their hobbies in the next three days. Then, in the last three days, each user is requested to repeat some searches with the queries entered in the previous days. We got a log of about 300 queries averaging 25 queries per subject and about 1200 records of the pages the users clicked in total. The size of this real data set is relatively small because the click data collection and users' judgments are labor intensive. The evaluation measure is MAP (mean average precision) which is widely used in ranking problems.

5. 2 Results and Discussions

In the real data set, the queries in the last three days are regarded as repeated IQs. The first seven-day click-through data is divided into two parts (odd-day and even-day) as semi-new informational searches. One is for setting up the QCW user profile and the other is for re-ranking search results based on the learned user profile, and then the two parts are exchanged to run the evaluation once again. Here, we set M to be 5 and $hf \cdot |U|$ to be 20.

In Figure 3 and Table 1, we summarize the performance of the proposed four re-rank strategies according to different hierarchical semantic measures. "MAP difference" means the difference value between our strategy and the baseline and "MAP%" represents the improvement percentage of our strategy over the baseline.



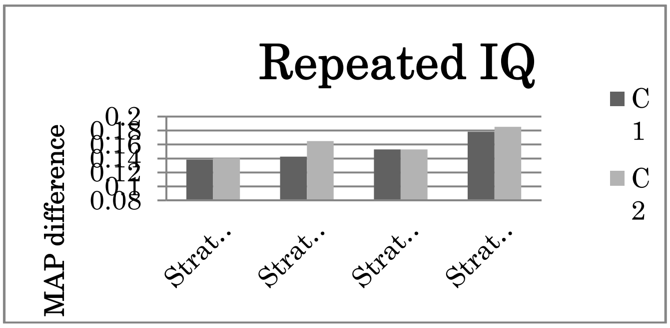


Figure 3 The MAP differences between our strategies and baseline

Table 1 The MAP improvement percentage of our strategies over baseline

Semi-new IQ(%)					Repeated IQ(%)			
Measure	Strategy 1	Strategy 2	Strategy 3	Strategy 4	Strategy 1	Strategy 2	Strategy 3	Strategy 4
C1	24.92	25.71	**27.13**	34.35	55.95	57.72	61.92	72.16
C2	**25.50**	**28.78**	26.64	**34.88**	**56.93**	**66.71**	**61.92**	**75.00**

The experimental results show that the proposed user-context aware re-rank strategies are more effective than the baseline. "Strategy 1", representing the general idea of most existing personalized re-ranking schemes, is inferior to other three strategies. Among the proposed four re-rank strategies, the "Strategy 4" broadly shows the best performance. The "Strategy 2" with selective utilization of user profiles, averagely produces better results than the "Strategy 1" and "Strategy 3". In Figure 3 and Table 1 the improvement of repeated IQs in is more obvious than those of semi-new IQs. The larger improvement of repeated IQs shows that our re-rank strategies can effectively retrieve the Web pages previously clicked by users since these queries have been submitted before and user's click behavior has been stored in our QCW.

Moreover, in Figure 3 and in Table 1 we observed that the similarity measures using nonlinear transformation function (i.e., C2 shown in orange columns) generally produce better performance that the similarity measures using linear transformation (i.e., C1 shown in blue columns). In a word, "Strategy 4" with C2 produces the largest improvement, e.g., its improvements over baseline are 34.88% and 75% for semi-new IQs and repeated IQs respectively.

From the results, we can say that re-ranking of search results through semantic based personalization actually can enhance the general search. We also confirm that there are two critical factors: (1) the query-to-query similarity which captures the long term search interests of a user (query dependent), and (2) the most recent search interest which reflects the short term search behavior of a user (time dependent). The two factors indicate that both short-term and long-term memories contribute to the improvement.

6. Related Work

In this section we give a brief overview of some related works in the literature of personalized search. There are two kinds of context information we can use to model search experience and capture user search histories. One is short-term context, which emphasizes that the most recent search is most directly close to the user's current information need [15], [25], [26]. Successive searches in a session usually have the same information need. Detecting a session boundary, however, is a difficult task. The other is long-term context, which generally assumes that users will hold their interests over a relatively long time. It means that any search in the past may have some effect on the current search [4], [14], [18], [30]. These studies commonly used all available contexts as a whole to improve the search result quality and ranking. Preliminary discussion on this problem in [29] is in the context of only exploiting long-term search history of users. In addition, several researchers have used taxonomic hierarchy (a simple directory based ontology) is used to represent user's interests in the Web search [4], [10], [16], [18], [20], [24]. However, very few have taken into consideration the hierarchical structure of the directory-based ontology when calculating similarity values between current search of a user and her search history. Chirita et al. [4] using hierarchical semantic measure, however, required users to manually select topics they are interested in. A unique characteristic of our STAR framework is the development of a selective use of personalized search history and a combination of long term and short term user search histories in rank optimization of personalized search.

7. Conclusions

We presented a STAR framework for selective utilization of user search behaviors for personalized learning and re-ranking. We designed a novel user search profile called query context window (QCW) to record the search behavior of a user. We developed a query-to-query similarity model and the fading memory based weight function. We showed how our STAR framework carefully chose and weighed the relevant click records as useful user context given an input query and how we applied hierarchical semantic similarity measures in our re-rank strategies. The experimental results show that our STAR approach to personalized search and re-ranking approach can effectively learn user-specific query-dependent personalization preference and significantly improve the accuracy of personalized search over the most existing personalized re-rankings. Our ongoing research includes designing an effective updating policy for user profiles, and more effective rank aggregation methods for further optimization of personalized search.

[References]

[1] M. S. Aktas, M. A. Nacar, and F. Menczer. Personalizing PageRank based on domain profiles. In WEBKDD, 83–90, 2004.
[2] D. Beeferman and A. L. Berger. Agglomerative clustering of a search engine query log. In KDD, 407–416, 2000.
[3] P.-A. Chirita, C. S. Firan, and W. Nejdl. Personalized query expansion for the web. In SIGIR, 7–14, 2007.
[4] P. A. Chirita, W. Nejdl, R. Paiu, and C. Kohlsch¨utter. Using ODP metadata to personalize search. In SIGIR, 178–185, 2005.
[5] Z. Dou, R. Song, and J.-R. Wen. A large-scale evaluation and analysis of personalized search strategies. In WWW, 581–590, 2007.
[6] L. Fitzpatrick and M. Dent. Automatic feedback using past queries: Social searching? In SIGIR, 306–313, 1997.
[7] N. S. Glance. Community search assistant. In IUI, 91–96, 2001.
[8] T. H. Haveliwala. Topic-sensitive pagerank: A context-sensitive ranking algorithm for web search. IEEE Trans. Knowl. Data Eng., 15(4):784–796, 2003.
[9] G. Jeh and J. Widom. Scaling personalized web search. In WWW, 271–279, 2003.
[10] H. R. Kim and P. K. Chan. Learning implicit user interest hierarchy for context in personalization. In IUI, 101–108, 2003.
[11] Y. Labrou and T. W. Finin. Yahoo! As an ontology: Using Yahoo! categories to describe documents. In CIKM, 180–187, 1999.
[12] L. Li, Z. Yang, B. Wang, and M. Kitsuregawa. Dynamic adaptation strategies for long-term and short-term user profile to personalize search. In APWeb/WAIM, 228–240, 2007.
[13] Y. Li, Z. Bandar, and D. McLean. An approach for measuring semantic similarity between words using multiple information sources. IEEE Trans. Knowl. Data Eng., 15(4):871–882, 2003.
[14] F. Liu, C. T. Yu, and W. Meng. Personalized web search for improving retrieval effectiveness. IEEE Trans. Knowl. Data Eng., 16(1):28–40, 2004.
[15] Y. Lv, L. Sun, J. Zhang, J.-Y. Nie, W. Chen, and W. Zhang. An iterative implicit feedback approach to personalized search. In ACL, 585–592, 2006.
[16] N. Nanas, V. S. Uren, and A. N. D. Roeck. Building and applying a concept hierarchy representation of a user profile. In SIGIR, 198–204, 2003.
[17] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford Digital Library Technologies Project, 1998.
[18] A. Pretschner and S. Gauch. Ontology based personalized search. In ICTAI, 391–398, 1999.
[19] F. Qiu and J. Cho. Automatic identification of user interest for personalized search. In WWW, 727–736, 2006.
[20] H. rae Kim and P. K. Chan. Personalized ranking of search results with learned user interest hierarchies from bookmarks. In WEBKDD, 32–43, 2005.
[21] V. V. Raghavan and H. Sever. On the reuse of past optimal queries. In SIGIR, 344–350, 1995.
[22] M. Richardson and P. Domingos. The intelligent surfer: Probabilistic combination of link and content information in pagerank. In NIPS, 1441–1448, 2001.
[23] D. E. Rose and D. Levinson. Understanding user goals in web search. In WWW, 13–19, 2004.
[24] V. Schickel-Zuber and B. Faltings. Inferring user's preferences using ontologies. In AAAI, 1413–1418, 2006.
[25] X. Shen, B. Tan, and C. Zhai. Context-sensitive information retrieval using implicit feedback. In SIGIR, 43–50, 2005.
[26] X. Shen, B. Tan, and C. Zhai. Implicit user modeling for personalized search. In CIKM, 824–831, 2005.
[27] M. Speretta and S. Gauch. Personalized search based on user search histories. In WI, 622–628, 2005.
[28] K. Sugiyama, K. Hatano, and M. Yoshikawa. Adaptive web search based on user profile constructed without any effort from users. In WWW, 675–684, 2004.
[29] B. Tan, X. Shen, and C. Zhai. Mining long-term search history to improve search accuracy. In KDD, 718–723, 2006.
[30] J. Teevan, S. T. Dumais, and E. Horvitz. Personalizing search via automated analysis of interests and activities. In SIGIR, 449–456, 2005.
[31] J.-R.Wen, J.-Y. Nie, and H. Zhang. Query clustering using user logs. ACM Trans. Inf. Syst., 20(1):59–81, 2002.
[32] J. Xu and W. B. Croft. Query expansion using local and global document analysis. In SIGIR, 4–11, 1996.
[33] J. Zobel and A.Moffat. Exploring the similarity space. SIGIR Forum, 32(1):18–34, 1998.

Lin LI
Ph.D student of Graduate School of Information Science and Technology, the University of Tokyo. She received the Master degree in Wuhan University of Technology, China. Her research interests include Web mining, personalization and recommendation.

Zhenglu YANG
Research associate of Institute of Industrial Science, the University of Tokyo. He received Ph.D degree in the University of Tokyo in 2008. His research interests include sequence mining and data mining.

Masaru KITSUREGAWA
Professor and the director of center for information fusion at Institute of Industrial Science, the University of Tokyo. He received the Ph.D degree in information engineering in 1983 from the University of Tokyo. His research interests include parallel processing and database engineering. He is a member of steering committee of IEEE ICDE, PAKDD and WAIM, and has been a trustee of the VLDB Endowment. He was the chair of data engineering special interest group of Institute of Electronic, Information, Communication, Engineering, Japan, the chair of ACM SIGMOD Japan, Chapter. He is currently a trustee of DBSJ.